Explicit Legg-Hutter intelligence calculations which suggest non-Archimedean intelligence

Samuel Allen Alexander[1][0000−0002−7930−110X]
Arthur Paul Pedersen[2][0000−0002−2164−6404]

[1]The U.S. Securities and Exchange Commission, samuelallenalexander@gmail.com
[2]The City University of New York

Abstract. Are the real numbers rich enough to measure intelligence? We generalize a result of Alexander and Hutter about the so-called Legg-Hutter intelligence measures of reinforcement learning agents. Using the generalized result, we exhibit a paradox: in one particular version of the Legg-Hutter intelligence measure, certain agents all have intelligence 0, even though in a certain sense some of them outperform others. We show that this paradox disappears if we vary the Legg-Hutter intelligence measure to be hyperreal-valued rather than real-valued.

1 Introduction

Legg and Hutter proposed [7] a theoretical measure of the intelligence of reinforcement learning (RL) agents—agents who interact with environments so as to maximize a reward signal. They proposed that the intelligence of an agent π be measured as $\Upsilon(\pi) = \sum_\mu 2^{-K(\mu)} V^\pi_\mu$, where μ ranges over the set of all suitably well-behaved computable RL environments, $K(\mu)$ is the Kolmogorov complexity of μ, and V^π_μ is the expected total reward π achieves in μ.

Because Kolmogorov complexity depends on the choice of a background universal Turing machine (UTM), the Legg-Hutter intelligence measure also implicitly depends on same. Alexander and Hutter showed [4] that if the UTM is symmetric in a certain sense, then agents satisfying a self-duality property have Legg-Hutter intelligence 0. In the present paper we generalize that result. We show that the Legg-Hutter intelligence of a self-dual agent π is $\sum_\nu 2^{-K(\nu)} V^\pi_\nu$ where ν is restricted to only range over those suitably well-behaved computable environments where the UTM is non-symmetric.

Armed with the above result, we will argue that for a certain choice of UTM, there are agents π_1 and π_2 such that $\Upsilon(\pi_1) = \Upsilon(\pi_2) = 0$ even though in some sense π_1 strictly outperforms π_2. We opine that this paradox is due to the inadequacy of the real numbers, $\mathbb{R}$, for measuring intelligence. Assuming a technical condition on the environments in the definition of Υ, we exhibit a variation Υ^* taking values in the hyperreal numbers (we do not assume prior familiarity with the hyperreals, so we gently introduce them using intuition about elections). The hyperreal-valued Legg-Hutter intelligence measure avoids the above-mentioned paradox.

The structure of the paper is as follows:

- In Section 2 we will informally describe our results.
- In Section 3 we will develop preliminaries.
- In Section 4 we will generalize a result of Alexander and Hutter.
- In Section 5 we will discuss what we call almost-symmetric universal Turing machines.
- In Section 6 we introduce what we call the Garden-of-Eden paradox.
- In Section 7 we introduce a variation of the Legg-Hutter intelligence measure which is hyperreal-number-valued instead of real-number-valued, and show that it solves the Garden-of-Eden paradox.
- In Section 8 we respond to anticipated objections.
- In Section 9 we summarize and make concluding remarks.

2 Informal Description of Results

Consider a "Garden of Eden" reinforcement learning environment in which one action is "forbidden" and all others are "allowed". If an agent takes the forbidden action even one time, then the agent's total reward from the environment becomes 0 (there is no way for the agent to recover from its sin). But if the agent never takes the forbidden action, then its total reward from the environment is 1.

Consider two different agents. Agent $A_{1\%}$ is an agent who, every turn, takes the forbidden action with 1% probability, or takes an allowed action with 99% probability. Agent $A_{99\%}$ is an agent who, every turn, takes the forbidden action with 99% probability, or takes an allowed action with 1% probability. Using standard real-valued probability theory, both agents have the same expected total reward in the Garden of Eden environment, namely, 0. This perhaps counterintuitive result is because, over infinitely many turns, the expected probability of $A_{1\%}$ eventually taking the forbidden action is 100%. There are non-standard variations of probability theory where $A_{1\%}$ has greater total expected reward than $A_{99\%}$ (but both these total expected rewards are infinitesimal—hence the necessity for the number system to be non-Archimedean, i.e., to fail to satisfy the Archimedean property of $\mathbb{R}$).

We show that by choosing the background universal Turing machine extremely carefully, we can arrange that for certain agents (including $A_{1\%}$ and $A_{99\%}$), the Legg-Hutter intelligence measure is entirely determined by performance in the Garden of Eden: because of said choice of universal Turing machine, the contributions from all other environments all perfectly cancel each other out. Thus, the above scandal is elevated from a paradox about probability theory to a paradox about intelligence measurement. Namely: $A_{1\%}$ and $A_{99\%}$ both have Legg-Hutter intelligence 0, even though in an intuitive sense $A_{1\%}$ is clearly the better agent if only the Garden of Eden environment matters. We present a variation of the Legg-Hutter intelligence measure, taking values in the hyperreal numbers (a non-Archimedean number system), where the paradox disappears.

3 Preliminaries

Fix a finite nonempty set $\mathcal{A}$ of *actions*, a finite nonempty set $\mathcal{O}$ of *observations*, and a finite nonempty set $\mathcal{R} \subseteq \mathbb{Q} \cap [-1, 1]$ of *rewards*. We assume that $\mathcal{R}$ is symmetric, in the sense that $\mathcal{R}$ contains $-r$ whenever $\mathcal{R}$ contains r. We also assume $0 \in \mathcal{R}$. We assume $\mathcal{A} \cap \mathcal{O} = \mathcal{A} \cap \mathcal{R} = \mathcal{O} \cap \mathcal{R} = \emptyset$. Let $\langle\rangle$ be the empty sequence. For any finite sequences s and t, let $s \frown t$ be the result of concatenating t to the end of s.

In the following definitions, we follow [4] except where otherwise indicated.

Definition 1. *(Reinforcement learning agents and environments)*

1. *Let $(\mathcal{ORA})^*$ be the set of finite sequences of the form $o_0, r_0, a_0, \ldots, o_k, r_k, a_k$ with each $o_i \in \mathcal{O}$, $r_i \in \mathcal{R}$, and $a_i \in \mathcal{A}$; we also include $\langle\rangle$ in $(\mathcal{ORA})^*$.*
2. *Let $(\mathcal{ORA})^*\mathcal{OR}$ be the set of all sequences of the form $s \frown \langle o, r\rangle$ where $s \in (\mathcal{ORA})^*$, $o \in \mathcal{O}$, $r \in \mathcal{R}$.*
3. *An* agent *is a function π, with domain $(\mathcal{ORA})^*\mathcal{OR}$, and with range the set of $\mathbb{Q}$-valued probability distributions on $\mathcal{A}$. For each $s \in (\mathcal{ORA})^*\mathcal{OR}$, we write $\pi(\bullet|s)$ for the value of π at s (so $\pi(\bullet|s)$ is a $\mathbb{Q}$-valued probability distribution on $\mathcal{A}$), and for each $a \in \mathcal{A}$, we write $\pi(a|s)$ for $(\pi(\bullet|s))(a)$; we think of $\pi(a|s)$ as the probability that the agent will take action a in response to stimulus s.*
4. *An* environment *is a function μ, with domain $(\mathcal{ORA})^*$ and with range the set of $\mathbb{Q}$-valued probability distributions on $\mathcal{O} \times \mathcal{R}$. For each $s \in (\mathcal{ORA})^*$, we write $\mu(\bullet|s)$ for the value of μ at s (so $\mu(\bullet|s)$ is a $\mathbb{Q}$-valued probability distribution on $\mathcal{O} \times \mathcal{R}$), and for each $(o, r) \in \mathcal{O} \times \mathcal{R}$, we write $\mu(o, r|s)$ for $(\mu(\bullet|s))(o, r)$; we think of $\mu(o, r|s)$ as the probability that the environment will issue observation o and reward r in response to stimulus s.*

Definition 2. *(Agent-environment interaction) Let π be any agent. Let μ be any environment.*

1. *For every $n \in \mathbb{N}$, let $V^\pi_{\mu,n}$ be the expected value of the sum of the rewards in the sequence $(o_0, r_0, a_0, \ldots, o_n, r_n, a_n)$ randomly generated as follows:*
   - *Choose $(o_0, r_0) \in \mathcal{O} \times \mathcal{R}$ randomly based on the probability distribution $\mu(\bullet|\langle\rangle)$.*
   - *Choose $a_0 \in \mathcal{A}$ randomly based on the probability distribution $\pi(\bullet|o_0, r_0)$.*
   - *For each $0 < k \leq n$, choose $(o_k, r_k) \in \mathcal{O} \times \mathcal{R}$ randomly based on the probability distribution $\mu(\bullet|o_0, r_0, a_0, \ldots, o_{k-1}, r_{k-1}, a_{k-1})$.*
   - *For each $0 < k \leq n$, choose $a_k \in \mathcal{A}$ randomly based on the probability distribution $\pi(\bullet|o_0, r_0, a_0, \ldots, o_{k-1}, r_{k-1}, a_{k-1}, o_k, r_k)$.*
2. *Let $V^\pi_\mu = \lim_{n\to\infty} V^\pi_{\mu,n}$ (if the limit converges to a real number; otherwise V^π_μ is undefined).*

Definition 3. *An environment μ is* well-behaved *if:*

- *μ is computable, and*
- *For every agent π, V^π_μ is defined and $-1 \leq V^\pi_\mu \leq 1$.*

Let W be the set of all well-behaved environments.

Definition 4. *(Duality)*

1. *For every sequence s, let $\overline{s}$ be the result of multiplying every reward in s by -1.*
2. *For every agent π, let $\overline{\pi}$, the* dual *of π, be the agent defined by $\overline{\pi}(a|s) = \pi(a|\overline{s})$ for all $a \in \mathcal{A}$, $s \in (\mathcal{ORA})^*\mathcal{OR}$.*
3. *For every environment μ, let $\overline{\mu}$, the* dual *of μ, be the environment defined by $\overline{\mu}(o, r|s) = \mu(o, -r|\overline{s})$ for all $o \in \mathcal{O}$, $r \in \mathcal{R}$, $s \in (\mathcal{ORA})^*$.*
4. *An agent π is* self-dual *if $\pi = \overline{\pi}$.*
5. *An environment μ is* self-dual *if $\mu = \overline{\mu}$.*

Lemma 1. *Let π be an agent, μ an environment, $n \in \mathbb{N}$.*

1. *$\overline{\overline{\pi}} = \pi$ and $\overline{\overline{\mu}} = \mu$.*
2. *$V^{\overline{\pi}}_{\overline{\mu},n} = -V^{\pi}_{\mu,n}$.*
3. *$V^{\overline{\pi}}_{\overline{\mu}} = -V^{\pi}_{\mu}$ (and the left-hand side is defined iff the right-hand side is).*
4. *$V^{\pi}_{\overline{\mu},n} = -V^{\overline{\pi}}_{\mu,n}$.*
5. *$V^{\pi}_{\overline{\mu}} = -V^{\overline{\pi}}_{\mu}$ (and the left-hand side is defined iff the right-hand side is).*
6. *μ is well-behaved iff $\overline{\mu}$ is well-behaved.*

Proof. Parts 1, 3, 5 and 6 are proved in [4]. The proofs of 3 and 5 in [4] work by proving 2 and 4. □

For any sets X and Y, we write $f : \subseteq X \to Y$ to indicate that f is a function with codomain Y and with domain some subset of X.

Definition 5. *(Prefix-free universal Turing machines) Let 2^* be the set of finite binary strings.*

1. *A function $f : \subseteq 2^* \to 2^*$ is* prefix-free *if f is computable and for all $p, p' \in 2^*$, if $f(p)$ and $f(p')$ are defined, then p is not a strict initial segment of p'.*
2. *A* prefix-free universal Turing machine *(or* PFUTM*) is a prefix-free function $U : \subseteq 2^* \to 2^*$ such that for every prefix-free $f : \subseteq 2^* \to 2^*$, $\exists y \in 2^*$ such that $\forall x \in 2^*$, $f(x) = U(y \frown x)$ (we call such a y a* computer program for f in programming language U*).*

We fix a computable Gödel numbering $\ulcorner \bullet \urcorner : (\mathcal{ORA})^* \to 2^*$ assigning to each $s \in (\mathcal{ORA})^*$ a code $\ulcorner s \urcorner \in 2^*$. Likewise, if M is the set of $\mathbb{Q}$-valued probability distributions on $\mathcal{O} \times \mathcal{R}$, we fix a computable Gödel numbering[1] $\ulcorner \bullet \urcorner : M \to 2^*$ assigning to each $p \in M$ a code $\ulcorner p \urcorner \in 2^*$. We assume that for all $x, y \in (\mathcal{ORA})^* \cup M$ with $x \neq y$, $\ulcorner x \urcorner$ is not an initial segment of $\ulcorner y \urcorner$ and $\ulcorner x \urcorner$ is not a terminal segment of $\ulcorner y \urcorner$. In other words, we assume $\ulcorner \bullet \urcorner$ is both prefix-free and suffix-free. By fixing $\ulcorner \bullet \urcorner$ we differ from [4], where the dependence of concepts such as Kolmogorov complexity (in the following definition) on $\ulcorner \bullet \urcorner$ was emphasized.

[1] This makes sense because $\mathcal{O} \times \mathcal{R}$ is finite.

Definition 6. *(Kolmogorov complexity)*

- *If* μ *is an environment and* U *is a PFUTM, we say that a function* $f : \subseteq 2^* \to 2^*$ encodes μ *if the following condition holds: For every* $s \in (\mathcal{ORA})^*$, $f(\ulcorner s \urcorner) = \ulcorner \mu(\bullet|s) \urcorner$.
- *For every computable environment* μ *and PFUTM* U*, let* $K_U(\mu)$ *be the* Kolmogorov complexity of μ given by U*, by which we mean the least* $n \in \mathbb{N}$ *such that there exists a computer program of length* n *in programming language* U *for some* $f : \subseteq 2^* \to 2^*$ *which encodes* μ.

Definition 7. *A PFUTM* U *is* symmetric *if the following condition holds: for every computable environment* μ, $K_U(\mu) = K_U(\overline{\mu})$.

In [4] it is shown that symmetric PFUTMs exist. In fact, the proof there shows more: there is a mechanical procedure for transforming any given PFUTM into a symmetric PFUTM.

In the following definition, we generalize the definition from [4] by introducing a new parameter W_0, for reasons which will become clear in Section 7.2. In [4], implicitly $W_0 = W$.

Definition 8. *(Legg-Hutter intelligence) Let* U *be a PFUTM and* $W_0 \subseteq W$ *a set of well-behaved environments. For every agent* π*, the* Legg-Hutter intelligence of π according to U, W_0 *is defined to be*

$$\Upsilon_{U,W_0}(\pi) = \sum_{\mu \in W_0} 2^{-K_U(\mu)} V_\mu^\pi$$

(the infinite sum is absolutely convergent by comparison with the sum defining Chaitin's constant, thus the sum does not depend on the order in which W_0 *is enumerated).*

The Legg-Hutter intelligence $\Upsilon_{U,W_0}(\pi)$ of an agent π (according to U, W_0) is intended to measure π's performance by averaging π's expected performance over the environments in W_0, using Kolmogorov complexity to assign lower weight to more contrived environments.

Theorem 1. *Let* U *be a symmetric PFUTM and let* π *be an agent.*

1. $\Upsilon_{U,W}(\overline{\pi}) = -\Upsilon_{U,W}(\pi)$.
2. *If* π *is self-dual then* $\Upsilon_{U,W}(\pi) = 0$.

Proof. See [4]. □

4 A Generalization of Theorem 1 part 2

Definition 9. *A set* W_0 *of well-behaved environments is* symmetric *if the following condition holds: for every* $\mu \in W_0$, $\overline{\mu} \in W_0$.

Note that W itself is symmetric by Lemma 1 part 6.

Lemma 2. *If environment μ and agent π are both self-dual, then $V^\pi_\mu = 0$.*

Proof. To show $V^\pi_\mu = 0$, it suffices to show $V^\pi_\mu = -V^\pi_\mu$. Compute:

$$
\begin{aligned}
V^\pi_\mu &= V^{\overline{\overline{\pi}}}_{\overline{\overline{\mu}}} && \text{(Lemma 1 part 1)}\\
&= -V^{\overline{\pi}}_{\overline{\mu}} && \text{(Lemma 1 part 3)}\\
&= -V^\pi_\mu. && \text{(Self-duality)}
\end{aligned}
$$

□

The following theorem is a generalization of Theorem 1 part 2.

Theorem 2. *Let U be any PFUTM. Let $W_0 \subseteq W$ be symmetric. Let $Z = \{\mu \in W_0 : K_U(\mu) \neq K_U(\overline{\mu})\}$ be the set of asymmetries of K_U. For any self-dual agent π, $\Upsilon_{U,W_0}(\pi) = \sum_{\mu \in Z} 2^{-K_U(\mu)} V^\pi_\mu$.*

Proof. Since $\sum_{\mu\in W_0} 2^{-K_U(\mu)} V^\pi_\mu = \sum_{\mu \in Z} 2^{-K_U(\mu)} V^\pi_\mu + \sum_{\mu \in W_0 \setminus Z} 2^{-K_U(\mu)} V^\pi_\mu$, it suffices to show $\sum_{\mu\in W_0\setminus Z} 2^{-K_U(\mu)} V^\pi_\mu = 0$. Let $W_1 = \{\mu \in W_0 \setminus Z : \mu \text{ is self-dual}\}$. Let $W_2 \subseteq W_0 \setminus (Z \cup W_1)$ be a maximal set such that for every $\mu \in W_2$, $\overline{\mu} \notin W_2$. Let $W_3 = \{\overline{\mu} : \mu \in W_2\}$. It follows that $W_0 \setminus Z$ is the disjoint union of W_1, W_2, W_3. Thus:

$$
\begin{aligned}
&\sum_{\mu \in W_0 \setminus Z} 2^{-K_U(\mu)} V^\pi_\mu\\
&= \sum_{\mu \in W_1} 2^{-K_U(\mu)} V^\pi_\mu + \sum_{\mu \in W_2 \cup W_3} 2^{-K_U(\mu)} V^\pi_\mu && (W_1 \cap (W_2 \cup W_3) = \emptyset)\\
&= \sum_{\mu \in W_1} 0 + \sum_{\mu \in W_2 \cup W_3} 2^{-K_U(\mu)} V^\pi_\mu && \text{(Lemma 2)}\\
&= \sum_{\mu \in W_2} (2^{-K_U(\mu)} V^\pi_\mu + 2^{-K_U(\overline{\mu})} V^\pi_{\overline{\mu}}) && \text{(Definition of } W_3)\\
&= \sum_{\mu \in W_2} (2^{-K_U(\mu)} V^\pi_\mu + 2^{-K_U(\mu)} V^\pi_{\overline{\mu}}) && (K_U(\mu) = K_U(\overline{\mu}) \text{ by def. of } Z)\\
&= \sum_{\mu \in W_2} (2^{-K_U(\mu)} V^\pi_\mu - 2^{-K_U(\mu)} V^{\overline{\pi}}_\mu) && \text{(Lemma 1 part 5)}\\
&= \sum_{\mu \in W_2} (2^{-K_U(\mu)} V^\pi_\mu - 2^{-K_U(\mu)} V^\pi_\mu) && (\pi \text{ is self-dual})\\
&= \sum_{\mu \in W_2} 0 = 0.
\end{aligned}
$$

□

5 Almost-symmetric PFUTMs

The infinite sum defining Legg-Hutter intelligence seems inherently intractable at first glance (if W_0 is nontrivial). Theorem 1 (part 2) was (in [4]) the first explicit computation of Legg-Hutter intelligence for a computable agent. In this section, we will develop machinery which will allow explicit computation of some non-integer Legg-Hutter intelligences.

Definition 10. *Let U be a PFUTM and let μ be a computable environment. We say U is* almost symmetric except at μ *if the following requirements are satisfied:*

1. $K_U(\mu) \neq K_U(\overline{\mu})$.
2. $K_U(\nu) = K_U(\overline{\nu})$ *for every computable environment* $\nu \notin \{\mu, \overline{\mu}\}$.

Proposition 1. *For every computable environment μ such that μ is not self-dual, there exists a PFUTM which is almost symmetric except at μ. In fact, for all positive integers $m \neq n$, there exists a PFUTM U such that $K_U(\mu) = m$, $K_U(\overline{\mu}) = n$, and $K_U(\nu) = K_U(\overline{\nu})$ for every computable environment $\nu \notin \{\mu, \overline{\mu}\}$.*

Proof. Let U_0 be any symmetric PFUTM. Let y_μ (resp. $y_{\overline{\mu}}$) be a computer program for μ (resp. $\overline{\mu}$) in programming language U_0. Let $t_1 = \langle 1 \rangle$, $t_2 = \langle 0, 1 \rangle$, $t_3 = \langle 0, 0, 1 \rangle$, and so on. Let $U :\subseteq 2^* \to 2^*$ be defined by

$$U(s) = \begin{cases} U_0(y_\mu \frown x) & \text{if } s = t_m \frown x; \\ U_0(y_{\overline{\mu}} \frown x) & \text{if } s = t_n \frown x; \\ U_0(x) & \text{if } s = t_{m+n} \frown x; \\ \text{undefined} & \text{in any other case.} \end{cases}$$

It is easy to see: U is a PFUTM; t_m is a computer program for μ in programming language U (and no shorter computer program for μ in programming language U exists); t_n is a computer program for $\overline{\mu}$ in programming language U (and no shorter computer program for $\overline{\mu}$ in programming language U exists); and that for every computable environment $\nu \notin \{\mu, \overline{\mu}\}$, the computer programs for ν in programming language U are exactly those strings $t_{m+n} \frown y$ such that y is a computer program for ν in programming language U_0. It follows that U witnesses the proposition. □

Theorem 3. *If μ is a well-behaved environment, U is a PFUTM which is almost-symmetric except at μ, $W_0 \subseteq W$ is symmetric, $\mu \in W_0$, and π is a self-dual agent, then*

$$\Upsilon_{U,W_0}(\pi) = (2^{-K_U(\mu)} - 2^{-K_U(\overline{\mu})}) V^\pi_\mu.$$

Proof. Compute:

$$\begin{aligned} \Upsilon_{U,W_0}(\pi) &= 2^{-K_U(\mu)} V^\pi_\mu + 2^{-K_U(\overline{\mu})} V^\pi_{\overline{\mu}} && \text{(Theorem 2)} \\ &= 2^{-K_U(\mu)} V^\pi_\mu - 2^{-K_U(\overline{\mu})} V^{\overline{\pi}}_\mu && \text{(Lemma 1 part 5)} \\ &= 2^{-K_U(\mu)} V^\pi_\mu - 2^{-K_U(\overline{\mu})} V^\pi_\mu. && \text{(Self-duality of } \pi) \end{aligned}$$

□

By carefully choosing μ and π, one can use Theorem 3 to obtain explicit nonzero Legg-Hutter intelligences. For example, for any nonzero $\alpha \in \mathcal{R}$, let μ be a computable environment which ignores the agent's actions and always gives initial reward α and reward 0 forever thereafter, regardless of anything the agent does. So $V^\pi_\mu = \alpha$ for every agent π. Since $\alpha \neq 0$, μ is not self-dual. By Proposition 1, there is, for example, a PFUTM U which is almost-symmetric except at μ, such that $K_U(\mu) = 1$, $K_U(\overline{\mu}) = 2$. By Theorem 3, $\Upsilon_{U,W}(\pi) = (2^{-1} - 2^{-2})V^\pi_\mu = \alpha/4$ for every self-dual agent π.

6 A Garden-of-Eden Paradox

For the rest of the paper, we assume there are at least two distinct actions in the action-set $\mathcal{A}$. We also assume that the reward-set $\mathcal{R}$ contains 1 and -1.

Definition 11. *Let $X \in \mathcal{A}$. An X-*forbidding Garden of Eden *is an environment μ such that the following conditions hold:*

1. *("μ gives initial reward 1 with probability 100%") For all $o \in \mathcal{O}$ and $r \in \mathcal{R}$, if $\mu(o, r|\langle\rangle) > 0$ then $r = 1$.*
2. *("μ gives reward -1 after the first X action, if ever, and reward 0 in all other situations") For every sequence $s \frown \langle a, o, r\rangle \in (\mathcal{ORA})^*\mathcal{OR}$ such that $\mu(o, r|s \frown \langle a\rangle) > 0$,*
   (a) If $a = X$ and X does not occur in s, then $r = -1$.
   (b) Otherwise, $r = 0$.

Thus, when an agent π interacts with an X-forbidding Garden of Eden μ, initially π gets a reward of 1. As long as π does not take action X, π maintains that reward of 1. But if π ever takes action X, then π immediately loses that full reward, and π's total cumulative reward is always 0 forever thereafter.

Definition 12. *For all $X, Y \in \mathcal{A}$ (with $X \neq Y$) and all $q \in [0, 1] \cap \mathbb{Q}$, let $\pi_{q,X,Y}$ be the agent which always takes action X with probability q or takes action Y with probability $1 - q$. So for every $c \in \mathcal{A}$ and every $s \in (\mathcal{ORA})^*\mathcal{OR}$,*

$$\pi_{q,X,Y}(c|s) = \begin{cases} q & \text{if } c = X \\ 1 - q & \text{if } c = Y \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 3. *For all q, X, Y as in Definition 12, $\pi_{q,X,Y}$ is self-dual.*

Proof. Clearly $\pi_{q,X,Y}(c|s)$ does not depend on the rewards in s, so $\pi_{q,X,Y}(c|s) = \pi_{q,X,Y}(c|\overline{s})$. By arbitrariness of s, $\overline{\pi_{q,X,Y}} = \pi_{q,X,Y}$. □

Lemma 4. *For each $X \in \mathcal{A}$, if μ is an X-forbidding Garden of Eden, then μ is not self-dual.*

Proof. Clearly $\overline{\mu}$ gives initial reward -1, whereas μ gives initial reward 1. Thus $\overline{\mu} \neq \mu$. □

Corollary 1. *For every $X \in \mathcal{A}$, there exists an X-forbidding Garden of Eden μ and a PFUTM U such that U is almost symmetric except at μ, and such that $K_U(\mu) = 1$ and $K_U(\overline{\mu}) = 2$.*

Proof. Clearly there exists some X-forbidding Garden of Eden μ. By Lemma 4, μ is not self-dual. The corollary now follows by Proposition 1. □

Theorem 4. *Let $X, Y \in \mathcal{A}$, with $X \neq Y$. Let μ be an X-forbidding Garden of Eden. Let U be a PFUTM which is almost symmetric except at μ, such that $K_U(\mu) = 1$ and $K_U(\overline{\mu}) = 2$. Let $W_0 \subseteq W$ be symmetric with $\mu \in W_0$. For every $q \in (0, 1] \cap \mathbb{Q}$, $\Upsilon_{U,W_0}(\pi_{q,X,Y}) = 0$.*

Proof. By Theorem 3, $\Upsilon_{U,W_0}(\pi_{q,X,Y}) = (2^{-1} - 2^{-2})V_\mu^{\pi_{q,X,Y}} = \frac{1}{4}V_\mu^{\pi_{q,X,Y}}$. For every $n \in \mathbb{N}$, if $(o_0, r_0, a_0, \ldots, o_n, r_n, a_n)$ are chosen randomly as in the definition of $V_{\mu,n}^{\pi_{q,X,Y}}$ (Definition 2), then, since $\pi_{q,X,Y}$ always takes action X with probability q, the probability that every $a_i \neq X$ (for $i = 0, \ldots, n-1$) is $(1-q)^n$. If so, then by Definition 11 it follows that $r_0 + \cdots + r_n = 1 + 0 + \cdots + 0 = 1$. Otherwise, by Definition 11 it follows that $r_0 + \cdots + r_n = 0$. Thus

$$V_{\mu,n}^{\pi_{q,X,Y}} = 1 \cdot (1-q)^n + 0 \cdot (1 - (1-q)^n).$$

Since $q \in (0, 1]$, it follows that $V_\mu^{\pi_{q,X,Y}} = \lim_{n\to\infty} V_{\mu,n}^{\pi_{q,X,Y}} = 0$. □

Theorem 4 is paradoxical because if $0 < q_1 < q_2 < 1$ then $\pi_{q_1,X,Y}$ ought to be strictly more performant than $\pi_{q_2,X,Y}$ in the X-forbidding Garden of Eden μ. If U, W_0 are as in Theorem 4 then Υ_{U,W_0} measures self-dual agent intelligence purely based on performance in μ. Thus, Υ_{U,W_0} measures intelligence entirely based on an agent's tendency to avoid taking the forbidden action X. Since $\pi_{q_1,X,Y}$ is less likely to take action X than $\pi_{q_2,X,Y}$ at each particular moment, the former agent ought to be considered more intelligent if intelligence is measured purely in terms of performance in this μ.

Theorem 4 shows that Legg-Hutter intelligence can be misleading even in a practical sense. Suppose we need an agent to perform in an X-forbidding Garden of Eden not for eternity, but for some unspecified positive number of steps. If our only options are $\pi_{q_1,X,Y}$ and $\pi_{q_2,X,Y}$, where $0 < q_1 < q_2 < 1$, then $\pi_{q_1,X,Y}$ is objectively the better choice, but $\Upsilon_{U,W}(\pi_{q_1,X,Y}) = \Upsilon_{U,W}(\pi_{q_2,X,Y}) = 0$ suggests that either option is just as good as the other.

This paradox is, of course, not surprising to the reader familiar with probability or measure theory. It falls under the same umbrella as the fact that if S and T are two countable subsets of $\mathbb{R}$ and m is, e.g., Lebesgue measure, then $m(S) = m(T) = 0$ even if S is a strict subset of T. In the Lebesgue measure case, what the paradox really shows is that m does not perfectly capture the notion of the size of a set. If it did, then $S \subsetneq T$ would imply $m(S) < m(T)$. In the same way, Theorem 4 shows that, at least for the contrived PFUTM in question, Legg-Hutter intelligence does not perfectly capture environmental performance of an agent.

This is not a condemnation of Legg-Hutter intelligence any more than it is a condemnation of Lebesgue measure. If the so-called *regularity* property, i.e. that

$S \subsetneq T$ implies $m(S) < m(T)$, is desired, one can attain it via measures taking their values from other number systems than the reals, such as the hyperreal number system; see [5]. In the next section, we will resolve the above Garden of Eden paradox, by introducing a hyperreal-valued variation of Legg-Hutter intelligence.

7 Legg-Hutter intelligence using nonstandard analysis

We will propose a hyperreal-valued variation of Legg-Hutter intelligence where the paradox in the previous section disappears. We do not assume the reader is familiar with the hyperreals, so we will briefly review (one construction of) the hyperreals.

7.1 Free ultrafilters and the hyperreal numbers

To intuitively motivate ultrafilters[2], it is instructive to imagine that the natural numbers are *voters* who cast ballots in order to decide true-or-false questions about functions $f : \mathbb{N} \to \mathbb{R}$. For example, if the question is whether or not $f : \mathbb{N} \to \mathbb{R}$ is larger on average than $g : \mathbb{N} \to \mathbb{R}$, we could consider each $n \in \mathbb{N}$ to vote as follows:

- If $f(n) > g(n)$, then n votes that f is larger than g on average.
- If $f(n) \leq g(n)$, then n votes that f is not larger than g on average.

One way to decide the outcome of such elections would be to decide in advance which sets of voters are *majorities*. Having suitably decided this, the winning candidate would be whichever candidate has a majority of voters vote for it. What properties should a choice of majorities satisfy? Three axioms immediately come to mind:

Definition 13. *Let $p \subseteq \mathscr{P}(\mathbb{N})$ be a set of subsets of $\mathbb{N}$, thought of as* majorities*.*

- *(Properness) p satisfies the* Properness *axiom if $\emptyset \notin p$. (If no-one votes for you, you lose.)*
- *(Monotonicity) p satisfies the* Monotonicity *axiom if the following requirement holds. For every $X \in p$, for every $Y \subseteq \mathbb{N}$, if $Y \supseteq X$ then $Y \in p$. (More votes can't hurt.)*
- *(Maximality) p satisfies the* Maximality *axiom if the following requirement holds. For every $X \subseteq \mathbb{N}$, either $X \in p$ or $X^c = \{n \in \mathbb{N} : n \notin X\} \in p$. (The election must have a winner.)*

[2] For a humorous presentation of this intuition in the form of a Socratic dialog, see [1]. This electoral motivation of ultrafilters was first made explicit in [2], though the theoretical underpinnings appeared in [6]. For a more direct application of the same idea to RL intelligence measurement, without any reference to Kolmogorov complexity or computability, see [3].

A fourth axiom is counter-intuitive when one thinks of elections, but intuitive when one considers that the answers to questions about transitive properties should be transitive. For example: if the voters decide that f is larger on average than g, and also that g is larger on average than h, then the voters ought to decide that f is larger on average than h. If $X_{fg} = \{n \in \mathbb{N} : f(n) > g(n)\}$, $X_{gh} = \{n \in \mathbb{N} : g(n) > h(n)\}$, and $X_{fh} = \{n \in \mathbb{N} : f(n) > h(n)\}$, then by linearity of $<$, we have $X_{fh} \supseteq X_{fg} \cap X_{gh}$. Thus, given Monotonicity, a simple way to force our election decision to be so consistent is to impose the following axiom.

Definition 14. *Let $p \subseteq \mathscr{P}(\mathbb{N})$ be a set of subsets of $\mathbb{N}$, thought of as* majorities*.*

– *($\cap$-Closure) p satisfies the $\cap$-Closure axiom if the following requirement holds. For all $X, Y \in p$, $X \cap Y \in p$.*

One trivial way to realize all four of the above axioms is as follows: choose some $n_0 \in \mathbb{N}$ as a *dictator* and declare that whoever n_0 votes for, automatically wins. For example, this would amount to declaring that f is larger on average than g iff $f(n_0) > g(n_0)$. This is clearly a poor way to decide elections. Therefore, we propose the following axiom.

Definition 15. *Let $p \subseteq \mathscr{P}(\mathbb{N})$ be a set of subsets of $\mathbb{N}$, thought of as* majorities*.*

– *(Non-Dictatorship) p satisfies the* Non-Dictatorship *axiom if the following requirement holds. For every $n_0 \in \mathbb{N}$, $\{n_0\} \notin p$.*

Although the above five axioms seem concrete, by combining them together we actually arrive at a mathematical concept which, without the above motivation, would seem quite abstract.

Definition 16. *A set $p \subseteq \mathscr{P}(\mathbb{N})$ of subsets of $\mathbb{N}$ is an* ultrafilter on $\mathbb{N}$ *(or simply an* ultrafilter*) if p satisfies the Properness, Monotonicity, Maximality, and $\cap$-closure axioms. An ultrafilter is* free *if it also satisfies the Non-Dictatorship axiom.*

The following lemma is well-known and we state it without proof. We mention, however, that logicians have proven that this lemma cannot be proved constructively; all of its proofs are necessarily non-constructive.

Lemma 5. *There exists a free ultrafilter on $\mathbb{N}$.*

Lemma 5 allows us to decide elections. Namely: fix a free ultrafilter p on $\mathbb{N}$, and declare that whenever the naturals vote in an election between candidates c_1 and c_2, then candidate c_i wins the election iff $\{n \in \mathbb{N} : n \text{ votes for } c_i\} \in p$. Such an i exists by Maximality; the Properness and $\cap$-Closure axioms ensure i is unique.

For the rest of the paper, fix a free ultrafilter p on $\mathbb{N}$.

Definition 17. *If $f, g : \mathbb{N} \to \mathbb{R}$, declare $f \sim g$ iff f and g are equally large on average, as voted by $\mathbb{N}$, using p to decide the election. That is, $f \sim g$ iff $\{n \in \mathbb{N} : f(n) = g(n)\} \in p$. Clearly $\sim$ is an equivalence relation.*

- *For every $f : \mathbb{N} \to \mathbb{R}$, let $[f]$ be the $\sim$-equivalence class which contains f. Let ${}^*\mathbb{R} = \{[f] : f : \mathbb{N} \to \mathbb{R}\}$; we call ${}^*\mathbb{R}$ the set of* hyperreal numbers.
- *For all $f, g : \mathbb{N} \to \mathbb{R}$, we define $[f] + [g] = [f + g]$ and $[f] \cdot [g] = [f \cdot g]$.*
- *For all $f, g : \mathbb{N} \to \mathbb{R}$, we declare $[f] < [g]$ iff g is larger on average than f, as voted by $\mathbb{N}$, using p to decide the election. That is, $[f] < [g]$ iff $\{n \in \mathbb{N} : f(n) < g(n)\} \in p$.*

The following lemma is well-known and we state it without proof.

Lemma 6. *The operations in Definition 17 parts 2 and 3 are well-defined. The resulting structure $({}^*\mathbb{R}, +, \cdot, <)$ is an ordered field extension of $\mathbb{R}$ (we consider $\mathbb{R}$ to be embedded in ${}^*\mathbb{R}$ by identifying every $r \in \mathbb{R}$ with the equivalence class $[n \mapsto r]$ of the corresponding constant function).*

In the following lemma, we collect a few well-known facts which we state without proof. These are straightforward to prove and we invite the reader to try to prove them.

Lemma 7. *1. If $X \subseteq \mathbb{N}$ is finite, then $X \notin p$.*
2. Suppose $f, g : \mathbb{N} \to \mathbb{R}$. If $f(n) > g(n)$ for all $n \in \mathbb{N}$, then $[f] > [g]$.
3. Suppose $f : \mathbb{N} \to \mathbb{R}$. If $\lim_{n\to\infty} f(n)$ exists, then for every real $\epsilon > 0$, the difference $|[f] - \lim_{n\to\infty} f(n)| < \epsilon$. In other words, the distance between $[f]$ and $\lim_{n\to\infty} f(n)$ is zero or infinitesimal.

7.2 Hyperreal-valued Legg-Hutter intelligence

In Definition 8, the infinite sum $\sum_{\mu \in W_0} 2^{-K_U(\mu)} V^\pi_\mu$ does not depend on the order in which W_0 is enumerated, because the sum is absolutely convergent. Said absolute convergence is a consequence of the requirement $-1 \leq V^\pi_\mu \leq 1$ in the definition of well-behaved environments (Definition 3). In order to define a hyperreal-valued Legg-Hutter intelligence, we would like to instead consider sums $\sum_{\mu \in W_0} 2^{-K_U(\mu)} V^\pi_{\mu,n}$ for various $n \in \mathbb{N}$. Unfortunately, such sums are not necessarily convergent, much less absolutely convergent. This is because even though $-1 \leq V^\pi_\mu \leq 1$, there is, a priori, no bound at all on $V^\pi_{\mu,n}$. For this reason, we must restrict attention to even better-behaved environments (this is the reason why we generalized Legg-Hutter intelligence by introducing the additional parameter W_0).

Definition 18. *An environment μ is* strongly well-behaved *if the following requirements hold.*

1. *μ is well-behaved.*
2. *For every agent π and $n \in \mathbb{N}$, $-1 \leq V^\pi_{\mu,n} \leq 1$.*

Lemma 8. *The set of strongly well-behaved environments is symmetric.*

Proof. Follows by Lemma 1 part 4. □

The following variation of Legg-Hutter intelligence is hyperreal-valued instead of real-valued. This increased granularity will allow the measure to distinguish between the agents which the real-valued measure failed to distinguish in Section 6.

Definition 19. *(Hyperreal-valued Legg-Hutter intelligence) Let U be a PFUTM and let $W_0 \subseteq W$ be a set of strongly well-behaved environments. For every agent π, the* hyperreal-valued Legg-Hutter intelligence of π according to U, W_0 *(and, implicitly, p) is defined to be*

$$\Upsilon^*_{U,W_0}(\pi) = \left[n \in \mathbb{N} \mapsto \sum_{\mu \in W_0} 2^{-K_U(\mu)} V^{\pi}_{\mu,n} \right] \in {}^*\mathbb{R}$$

(the infinite sums in question are defined, and are independent of the order in which W_0 is enumerated, because they are absolutely convergent by comparison with Chaitin's constant).

We want to show that Υ^*_{U,W_0} does not differ much from Υ_{U,W_0}. In order to do this, we will need a theorem from real analysis called Tannery's Theorem. This theorem is standard, so we state it without proof.

Lemma 9. *(Tannery's Theorem) Assume $\{a_i : \mathbb{N} \to \mathbb{R}\}_{i=0}^{\infty}$ is a sequence of sequences such that each $\lim_{n\to\infty} a_i(n)$ converges. Assume $w_0, w_1, \ldots \in \mathbb{R}$ satisfy $\sum_{i=0}^{\infty} w_i < \infty$ and for all $i, n \in \mathbb{N}$, $|a_i(n)| \leq w_i$. Then*

$$\lim_{n\to\infty} \sum_{i=0}^{\infty} a_i(n) = \sum_{i=0}^{\infty} \lim_{n\to\infty} a_i(n).$$

The following theorem shows that $\Upsilon^*_{U,W_0}(\pi)$ is infinitely close to $\Upsilon_{U,W_0}(\pi)$.

Theorem 5. *For any U, W_0, π as in Definition 19, $\Upsilon_{U,W_0}(\pi)$ and $\Upsilon^*_{U,W_0}(\pi)$ differ by an amount smaller than any positive real number.*

Proof. We assume W_0 is infinite (the other case is similar and easier). Let $\mu_0, \mu_1, \ldots$ enumerate W_0 (this is possible because the environments in W_0 are well-behaved, thus computable, thus countable). By Tannery's Theorem (Lemma 9) with $a_i(n) = 2^{-K_U(\mu_i)} V^{\pi}_{\mu_i,n}$ and $w_i = 2^{-K_U(\mu_i)}$,

$$\lim_{n\to\infty} \sum_{i=0}^{\infty} 2^{-K_U(\mu_i)} V^{\pi}_{\mu_i,n} = \sum_{i=0}^{\infty} \lim_{n\to\infty} 2^{-K_U(\mu_i)} V^{\pi}_{\mu_i,n}.$$

Since all the $\mu \in W_0$ are strongly well-behaved, it follows that all the infinite sums in question are absolutely convergent and so do not depend on the order of summation, so we can conclude $(*)$

$$\lim_{n\to\infty} \sum_{\mu \in W_0} 2^{-K_U(\mu)} V^{\pi}_{\mu,n} = \sum_{\mu \in W_0} \lim_{n\to\infty} 2^{-K_U(\mu)} V^{\pi}_{\mu,n}.$$

Thus:

$$\begin{aligned}
\Upsilon_{U,W_0}(\pi) &= \sum_{\mu\in W_0} 2^{-K_U(\mu)} V^\pi_\mu && \text{(Definition 8)}\\
&= \sum_{\mu\in W_0} 2^{-K_U(\mu)} \lim_{n\to\infty} V^\pi_{\mu,n} && \text{(Definition 2 part 2)}\\
&= \sum_{\mu\in W_0} \lim_{n\to\infty} 2^{-K_U(\mu)} V^\pi_{\mu,n} && \text{(Algebra)}\\
&= \lim_{n\to\infty} \sum_{\mu\in W_0} 2^{-K_U(\mu)} V^\pi_{\mu,n}. && \text{(By } *\text{)}
\end{aligned}$$

The theorem now follows by Lemma 7 part 3. □

To show that Υ^*_{U,W_0} avoids the Garden of Eden paradox, we will need to restate some of our above results for finite values of n instead of $n = \infty$.

Lemma 10. *Let $n \in \mathbb{N}$.*

1. *(Compare Lemma 2) If π is a self-dual agent and μ is a self-dual environment, then $V^\pi_{\mu,n} = 0$.*
2. *(Compare Theorem 2) For any PFUTM U, for any symmetric $W_0 \subseteq W$, if $Z = \{\mu \in W_0 : K_U(\mu) \neq K_U(\overline{\mu})\}$ is the set of asymmetries of K_U, then for any self-dual agent π,*
$$\sum_{\mu\in W_0} 2^{-K_U(\mu)} V^\pi_{\mu,n} = \sum_{\mu\in Z} 2^{-K_U(\mu)} V^\pi_{\mu,n}.$$
3. *(Compare Theorem 3) For any strongly well-behaved environment μ, for any PFUTM U which is almost-symmetric except at μ, for any symmetric $W_0 \subseteq W$ with $\mu \in W_0$, for any self-dual agent π,*
$$\sum_{\nu\in W_0} 2^{-K_U(\nu)} V^\pi_{\nu,n} = (2^{-K_U(\mu)} - 2^{-K_U(\overline{\mu})}) V^\pi_{\mu,n}.$$

Proof. (1) Similar to the proof of Lemma 2, but use Lemma 1 part 2 instead of Lemma 1 part 3.

(2) Similar to the proof of Theorem 2, but use (1) instead of Lemma 2 and use Lemma 1 part 4 instead of Lemma 1 part 5.

(3) Similar to the proof of Theorem 3, but use (2) instead of Theorem 2 and use Lemma 1 part 4 instead of Lemma 1 part 5.

Finally, we show that the hyperreal Legg-Hutter intelligence measure is free of the pathological behavior from Section 6.

Theorem 6. *(Contrast Theorem 4) Let $X, Y \in \mathcal{A}$, with $X \neq Y$. Let μ be a well-behaved X-forbidding Garden of Eden. Let U be a PFUTM which is almost symmetric except at μ, such that $K_U(\mu) = 1$ and $K_U(\overline{\mu}) = 2$. Let W_0 be a symmetric set of strongly well-behaved environments, with $\mu \in W_0$. For all $q_1, q_2 \in (0,1] \cap \mathbb{Q}$ with $q_1 < q_2$, we have $\Upsilon^*_{U,W_0}(\pi_{q_1,X,Y}) > \Upsilon^*_{U,W_0}(\pi_{q_2,X,Y})$.*

Proof. By Definition 11, clearly μ is strongly well-behaved. By Lemma 10 part 3, for every $n \in \mathbb{N}$, for each $i \in \{1,2\}$, $\sum_{\nu \in W_0} 2^{-K_U(\nu)} V_{\nu,n}^{\pi_{q_i},X,Y} = \frac{1}{4} V_{\mu,n}^{\pi_{q_i},X,Y}$. As in the proof of Theorem 4, for every $n \in \mathbb{N}$, for each $i \in \{1,2\}$,

$$V_{\mu,n}^{\pi_{q_i},X,Y} = 1 \cdot (1-q_i)^n + 0 \cdot (1-(1-q_i)^n).$$

The theorem now follows by Lemma 7 part 2. □

There might be other approaches to Legg-Hutter intelligence avoiding the Garden of Eden paradox. For example, in [8], Pedersen builds alternate foundations of nonstandard probability theory along similar lines to de Finetti's foundations of standard probability, without certain limitations of the latter. It might be possible to apply these nonstandard probability theory foundations to the problem.

8 Anticipated Objections

8.1 What does it really matter if the agent takes the forbidden action only 1% of the time or 99% of the time? In an infinite Garden of Eden interaction, either agent will eventually take the forbidden action with probability 100%.

In standard probability theory, an event having probability 100% does not necessarily mean that it is certain. In order to resolve the probability distribution output by the agent on the nth turn into an actual action, we might imagine that a random number $x_n \in [0,1)$ is generated. For the 1% agent, the nth action is the forbidden action iff $x_n < 0.01$. For the 99% agent, the nth action is the forbidden action iff $x_n < 0.99$. If $S_{1\%} = \{(x_1, x_2, \ldots) \in [0,1)^\infty : \forall n,\, x_n \geq 0.01\}$ (the event of the 1% agent going for all eternity without taking the forbidden action), and $S_{99\%} = \{(x_1, x_2, \ldots) \in [0,1)^\infty : \forall n,\, x_n \geq 0.99\}$ (the event of the 99% agent going all eternity without taking the forbidden action), then $S_{1\%} \supsetneq S_{99\%}$.

Furthermore, in actual practice, we never run an RL agent for all eternity. At most, we run the agent for some indeterminite finite number of steps. Clearly the 1% forbidden action agent beats the 99% forbidden action agent in this case.

8.2 The PFUTM in Theorem 4 is too contrived for us to draw conclusions about intelligence measurement in more realistic contexts

We conjecture that similar paradoxes are embedded in Legg-Hutter intelligence measures based on more familiar PFUTMs, but it is difficult to explicitly exhibit them because of the intractible nature of the infinite sum defining Legg-Hutter intelligence.

8.3 The non-constructive nature of Lemma 5 renders Υ^*_{U,W_0} impractical to calculate

It is already impossible to compute Υ_{U,W_0}. In fact, the Kolmogorov complexity function itself is already non-computable, so we cannot generally even compute individual summands in the infinite sum defining $\Upsilon_{U,W_0}(\pi)$. One could actually argue that in a sense, Υ^*_{U,W_0} is *easier* to approximate than Υ_{U,W_0}. Here is what we mean by this. If one were to approximate the Legg-Hutter intelligence of $\pi_{q_1,X,Y}$ and $\pi_{q_2,X,Y}$ (in the context of Theorem 4) by running large amounts of finite agent-environment interactions (Monte Carlo style), one would see $\pi_{q_1,X,Y}$ outperforming $\pi_{q_2,X,Y}$ if $q_1 \ll q_2$, which is consistent with Theorem 6 and inconsistent with Theorem 4.

8.4 A better way to resolve the paradox would be to use discount factors

The Garden-of-Eden paradox in Theorem 4 would disappear if one applied a discount factor in the definition of V^π_μ, say, weighing each nth reward by γ^n for some fixed discount factor $\gamma \in (0,1)$. And indeed, one could treat the paradox as evidence in favor of applying such discount factors. But Legg and Hutter specifically elected, in [7], not to use discount factors, and gave good reasons for their decision.

9 Summary and conclusion

In Theorem 2 we generalized a result of Alexander and Hutter [4]. In Theorem 4 we used this to show that if the background universal Turing machine is carefully chosen, so that Legg-Hutter intelligence measures performance in one particular "Garden of Eden" environment, then, paradoxically, certain agents all have Legg-Hutter intelligence 0 despite the fact that in some sense some of them outperform others in said environment. We opine that this Garden-of-Eden paradox results from the coarseness of the real numbers. In Theorems 5 and 6 we show that the paradox can be resolved by allowing the Legg-Hutter intelligence measure to take its values from the hyperreal number system, a more granular number system than $\mathbb{R}$.

Acknowledgments

We gratefully acknowledge Aram Ebtakar and Cole Wyeth for comments and feedback.

References

1. Samuel Alexander and Bryan Dawson. Big-oh notations, elections, and hyperreal numbers: A Socratic dialogue. *Proceedings of the ACMS*, 23:15–22, 2022.

2. Samuel Alexander and Bill Hibbard. Measuring intelligence and growth rate: variations on Hibbard's intelligence measure. *Journal of Artificial General Intelligence*, 12(1):1–25, 2021.
3. Samuel Allen Alexander. Intelligence via ultrafilters: structural properties of some intelligence comparators of deterministic Legg-Hutter agents. *Journal of Artificial General Intelligence*, 10(1):24–45, 2019.
4. Samuel Allen Alexander and Marcus Hutter. Reward-punishment symmetric universal intelligence. In *Artificial General Intelligence: 14th International Conference*, pages 1–10. Springer, 2022.
5. Thomas Hofweber and Ralf Schindler. Hyperreal-valued probability measures approximating a real-valued measure. *Notre Dame Journal of Formal Logic*, 57, 2016.
6. Alan P Kirman and Dieter Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972.
7. Shane Legg and Marcus Hutter. Universal intelligence: A definition of machine intelligence. *Minds and machines*, 17:391–444, 2007.
8. Arthur Paul Pedersen. Comparative expectations. *Studia Logica*, 102:811–848, 2014.